

What if pupils can assess their peers anonymously?

What if pupils can assess their peers anonymously?

A quasi-experimental study.

Ellen Vanderhoven^{a,b}, Annelies Raes^a, Hannelore Montrieux^a, Tijs Rotsaert^a and Tammy Schellens^a

^a Department of Educational Studies, Ghent University

^b iMinds, MICT, Ghent University

annelies.raes@ugent.be, ellen.vanderhoven@ugent.be, hannelore.montrieux@ugent.be,
tijs.rotsaert@ugent.be, tammy.schellens@ugent.be

Correspondence concerning this manuscript should be addressed to:

Ellen Vanderhoven

Department of Educational Studies, Ghent University,

Henri Dunantlaan 2, 9000 Gent.

Tel.: +329 3310316

Fax: +329 3310320

E-mail: Ellen.Vanderhoven@ugent.be

Note: Manuscript accepted for publication in Computers & Education

What if pupils can assess their peers anonymously?

What if pupils can assess their peers anonymously? A quasi-experimental study.

Abstract

Peer assessment has proven to be a promising assessment form, but there is only limited research about how to decrease the undesirable social effects that are inherent to the process, such as peer pressure and fear of disapproval. In previous research, anonymity has proven to be an important factor in peer assessment in higher education. In the current research, peer assessment was studied in secondary education and classroom response technology (CRT) was introduced as a tool that enables anonymity within face-to-face settings. A quasi-experimental study was set up in four classes to compare traditional non-anonymous peer assessment (raising score cards) with anonymous peer assessment (giving scores using CRT). It was questioned whether students felt more positive towards anonymous peer assessment, and reported to feel less negative social effects. Subsequently, the hypothesis that anonymous peer assessment would be a more valid methodology is verified. Finally, teachers' experiences with both peer assessment interventions were studied. Although some concerns were raised about the validity of anonymous peer assessment, it has been found that pupils felt more positive towards peer assessment and experienced less peer pressure and fear of disapproval when scores were given anonymously using CRT. Teachers reported that using CRT was an objective way of assessing but raised some concerns with regard to the control of the teacher, and the classroom characteristics. In this regard, implications for future research are discussed.

Keywords: peer assessment, anonymity, computer-assisted learning, secondary education; evaluation methodologies; interactive learning environments; computer-mediated communication; cooperative/collaborative learning

What if pupils can assess their peers anonymously?

1. Introduction

The so-called 21st century skills are a list of skills that aim to prepare students for complex professional tasks in increasingly complex workplaces (Dumont & Istance, 2010; Scardamalia, 2001). To satisfy these 21st century needs, contemporary education aims at self-directed and collaborative learning with active participation of the learner (see e.g., Boud, Cohen, & Sampson, 1999). Following this evolution, the notion of assessment became important as well, as literature has indicated that student learning can be positively influenced by the manner in which assessment practices are introduced in the classroom (e.g., Biggs, 1996; Birenbaum, 2003; Black & Wiliam, 1998; Sluijsmans, 2002). Van Gennip, Segers, and Tillema (2010) state that assessment has the potential to inform students about their strengths and weaknesses and to indicate the next steps one should take in the learning process in order to perform better in subsequent performances.

In the context of this learning approach inspired by social constructivism, this evolution in evaluation entails different innovative forms of assessment like self, peer and co assessment (Fastré, van der Klink, Sluijsmans, & van Merriënboer, 2013; Harris & Brown, 2013; Raes, Vanderhoven, & Schellens, 2013). Yet, both teachers and students need to adjust to these new assessment forms and this is especially the case in assessment practices in which students get the responsibility to evaluate each other, such as peer assessment, and where interpersonal variables might influence the procedure (Van Gennip, Segers, & Tillema, 2009). Research about peer assessment is still in a stage of adolescence (Kollar & Fischer, 2010) and there is a growing interest in the impact of interpersonal variables on the outcome of peer assessment practices (Hovardas, Tsivitanidou, & Zacharia, 2014; Panadero, Romero, & Strijbos, 2013). While researchers suggest that anonymity might offer a solution for the possible negative impact of these interpersonal variables (Ainsworth et al., 2011), the use of anonymity in combination with assessment is still an understudied issue (Panadero et al.,

What if pupils can assess their peers anonymously?

2013). This study tries to meet these research gaps by setting up a quasi-experimental study in authentic secondary education. In this study it is investigated how undesirable social effects can be reduced by increasing the anonymity using classroom response technology during face-to-face peer assessment in secondary education.

1.1 Peer assessment and its social nature

Peer assessment is an educational setting that gives the opportunity for students to be actively involved in the assessment of their peers (Kollar & Fischer, 2010). This occurs by considering and specifying the level, value or quality of a product or performance of their peers, by means of oral and/or written feedback (Topping, 1998, 2003). Given certain conditions, such as the presence of unambiguous criteria on which to evaluate (Falchikov & Goldfinch, 2000), a training in peer assessment (Sluijsmans, Brand-Gruwel, van Merriënboer, & Martens, 2004; van Zundert, Sluijsmans, & van Merriënboer, 2010), and the use of rubrics (Panadero et al., 2013) this kind of evaluation has proven to be accurate, with high inter-rater agreements between peers' and teachers' scores (Falchikov & Goldfinch, 2000). In addition, it has been proven that pupils involved in peer assessment perform better, because they have a better understanding of the assessment criteria (Li, Liu, & Steckelberg, 2010; Smith, Cooper, & Lancaster, 2002; Topping, 2003).

However, it needs to be taken into account that peer assessment has a social nature since it occurs in an interactive setting, in which at least two peers need to collaborate (Kollar & Fischer, 2010; van Gennip et al., 2010). As mentioned before, the development and interplay between interpersonal variables, such as peer pressure and fear of disapproval when giving low scores, can therefore affect the outcomes of the assessment (Barron, 2003; Cartney, 2010; Panadero et al., 2013; van Gennip et al., 2009). These effects are twofold. First, negative feelings during peer assessment can be considered undesirable since they might undermine the motivation for participation during the procedure (Stepanyan, Mather, Jones, &

What if pupils can assess their peers anonymously?

Lusuardi, 2009). Second, researchers describe reciprocity effects, referring to the bias on the peer assessment outcome caused by interpersonal variables, potentially jeopardizing the validity of the assessment procedure (Strijbos & Sluijsmans, 2010).

According to the social impact theory of Latané (1981), the effects of interpersonal variables can be even stronger in face-to-face classroom settings, as it states that social influence such as peer pressure will increase when others are close-by. This is confirmed by several researchers who have found that students in face-to-face classroom settings do not feel comfortable and can experience stress when publicly evaluating their peers (Pope, 2005; Stepanyan, Mather, Jones, & Lusuardi, 2009).

1.2 Anonymity within peer assessment practices

As the interplay of interpersonal variables is inherent to peer assessment practices and this process influences the assessment outcome, it has been stated that decreasing negative social effects such as peer pressure is desirable (Harris & Brown, 2013; Raes et al., 2013; Sung, Chang, Chang, & Yu, 2010). Several theories predict that this decrease can be obtained by making the assessment procedure anonymous, enabling pupils to assess each other without revealing their identity (Ainsworth et al., 2011; Deutsch & Gerard, 1955; Latané, 1981). For example, the social impact theory (Latané, 1981) states that identifiability, as opposed to anonymity, would increase social influence and the theory of normative influence of Deutsch and Gerard (1955) predicts that normative social influence upon the individual judgment (e.g., peer pressure when assessing a peer) decreases when people can judge anonymously. In line with these theories, Howard, Barrett, and Frick (2010) found that students who were anonymous when giving feedback in the context of an asynchronous webforum were approximately five times more likely to provide substantively critical feedback than those whose identities were known to the recipients. Moreover, Raes et al. (2013) established that providing anonymity in a face-to-face peer assessment context in higher education is

What if pupils can assess their peers anonymously?

associated with some interpersonal variables, that is with increased feelings of comfort and more positive attitudes towards peer assessment.

1.3 Establishing anonymity by means of Classroom Response Technology

The advent of information technologies and the internet brings new possibilities for peer assessment practices, for example by using wikis within online learning management systems (Gielen & De Wever, 2012; Liu & Li, 2013). This has many advantages, such as the ease of anonymizing the participants by using identification numbers or pseudonyms, while data can be automated and summarized (Liu & Li, 2013). However, as opposed to an online learning environment, a face-to-face classroom setting allows for a more interactive variant of peer assessment, as immediate and synchronous feedback can be provided. Yet, in this context anonymous peer assessment is more complicated to organize (Ainsworth et al., 2011).

Again, new technology offers possibilities. Raes et al. (2013) proposed and investigated the use of classroom response technology (CRT, e.g., the electronic voting system TurningPoint®) to enable pupils to give immediate anonymous feedback even within face-to-face settings and to help students cope with the aforementioned undesirable social effects. A classroom response system is a voting system used in a face-to-face setting to assess students using individual infrared handset transmitters ('clickers'), tablets or mobile phones. The aggregated totals of votes are subsequently displayed on a screen in front of the classroom as immediate feedback.

An extensive literature review shows that CRT is an effective educational tool when used during courses in higher education (Kay & LeSage, 2009) and several studies show that it improves students' engagement and learning (e.g., Han & Finkelstein, 2013). Recent research on the use of clickers in collaborative learning settings in higher education shows that when interactivity is present, students will be more motivated, attentive and participative, and they will be more likely to exchange ideas with each other (Blasco-Arcas, Buil,

What if pupils can assess their peers anonymously?

Hernández-Ortega, & Sese, 2013). Moreover, students report that it is the anonymous nature of the response that encourages them to participate (Bojinova & Oigara, 2011; Brady, Seli, & Rosenthal, 2013; Draper & Brown, 2004). With regard to peer assessment, CRT can therefore be used to allow students to score their peers without revealing their own identity and to allow the assessees to get immediate feedback.

1.4 Gaps in peer assessment research

Although interpersonal variables might decrease the motivation for participation in peer assessment, and might jeopardize the validity of the assessment procedure, research that explores the role of anonymity within face-to-face peer assessment is scarce (Panadero et al., 2013). Moreover, the limited research in this context is situated in higher education (Raes et al., 2013). Yet, the authors of this research report that the impact of interpersonal variables might be even stronger in secondary education, since it has been found that teenagers in general perceive higher levels of peer pressure than young adults (Steinberg & Monahan, 2007; Sumter, Bokhorst, Steinberg, & Westenberg, 2009). However, there is a lack of (quasi-experimental) research about peer assessment in secondary education (Strijbos & Sluijsmans, 2010; van Gennip et al., 2009; van Zundert et al., 2010). For this reason, it is interesting to find out the impact of anonymous peer assessment using CRT in a face-to-face setting in secondary education, both on the presence of undesirable social effects and on the validity of the assessment procedure.

Additionally, Raes et al. (2013) raise the possibility that displaying immediate visual feedback as a graph on a screen in front of the whole class, might cause negative feelings with the person being evaluated. Since they did not measure feelings of comfort with the assessed students in their study, no conclusion could be drawn about this issue based on their data. Again, this discomfort might be even bigger for teenagers in secondary education, given the importance of the opinion of their peers (Sumter et al., 2009). Following the remark of van

What if pupils can assess their peers anonymously?

Zundert et al. (2010) concerning the lack of differentiation between giving and receiving feedback, it is therefore also interesting to explore the feelings of the teenagers being evaluated.

Finally, next to taking into account the feelings of the assessors and assessees, it is interesting to explore the feelings of the classroom teacher as well. While the role of the teacher is clearly changing from the pivotal grading authority towards a facilitator of the (peer) assessment procedure (Fastré et al., 2013), no research could be found about the teachers' emotions in this process.

2. Research questions and hypotheses

In this manuscript, we elaborate on the results of previous research and try to fill the aforementioned gaps revealed in the literature study in the context of anonymous peer assessment using CRT in secondary education. This quasi-experimental study added to the literature by examining the effect of a specific peer assessment mechanism in an ecologically valid research setting (Topping, 2010). A traditional non-anonymous peer assessment setting was compared with an anonymous peer assessment setting using CRT in a secondary classroom with regard to different social effects. Preliminary analyses showed promising results, that is that pupils felt more positive towards peer assessment in the anonymous setting, and that they felt less negative emotions such as peer pressure or fear of disapproval (Author, 2012)¹. However, an elaborated literature study revealed other gaps that were not tackled in previous analyses (see section 1.4). Further analyses were needed to fill all the research gaps, and to extend the understanding of the findings, based on both quantitative and qualitative data. In this respect, the following research questions and hypotheses were put forth:

¹ Preliminary results have been published in a conference paper (Author et al., 2012)

What if pupils can assess their peers anonymously?

- RQ1: What is the effect of providing anonymity during peer assessment on undesirable social effects?

Based on previous research presented above (e.g. Ainsworth et al., 2011; Deutsch & Gerard, 1955; Latané, 1981), concerning the perspective of the assessor, first it is hypothesized that anonymous peer assessment can result in a decrease of negative social effects, that is pupils' experienced peer pressure and fear of disapproval when evaluating (hypothesis 1a). Yet, on the other hand, from an assessee's point of view, it is secondly hypothesized that students can feel uncomfortable when receiving immediate visual feedback in front of the class.

- RQ2: What is the effect of providing anonymity on pupils' evaluations of peer assessment?

Building on the first hypothesis, second it is hypothesized that pupils will have more positive attitudes towards anonymous peer assessment since previous literature state that assessing each other without revealing their identity can result in a decrease of negative social effects and in turn leads to more positive attitudes towards peer assessment (Ainsworth et al., 2011; Deutsch & Gerard, 1955; Latané, 1981).

- RQ3: Is the anonymous peer assessment procedure using CRT a valid method for assessment?

Since it is assumed that providing anonymity decrease undesirable social effects which potentially jeopardize the validity of the assessment procedure (Strijbos & Sluijsmans, 2010), third it is hypothesized that the anonymous peer assessment procedure using CRT is a more valid method for assessment compared to the traditional peer assessment procedure.

- RQ4: What is the effect of providing anonymity during peer assessment on teachers' experiences of peer assessment?

What if pupils can assess their peers anonymously?

Since no previous research could be found about the teachers' emotions in this process, the nature of this research question is more explorative and no hypothesis has been put forth.

3. Method

3.1 Participants and context

Two teachers were involved in this study and each of them participated in the study with two of their secondary classes (grade 9 and 10). This resulted in 69 participating pupils aged between 15 and 16 years old, of which 72% were girls. In accordance with previous research (De Grez, Valcke, & Roozen, 2009; Raes et al., 2013), peer assessment has been applied to assess students' oral presentations skills. Based on the compulsory curriculum and attainment targets for these grades, the third author constructed the group assignment and rubrics that would be used for the peer assessment intervention together with the participating teachers. The pupils were assigned to groups of two or three, and were instructed to prepare a presentation of about 15 minutes. The topics selected for the presentations were "Carolingian art" or "Mythical personalities of the 21st century". Both teachers and their students had no prior experiences with peer assessment.

3.2 Design & procedure

Since this study aimed to investigate the effect of anonymity in peer assessment through a quasi-experimental design, classes were randomly assigned to one of two conditions, so that both teachers had one class in the experimental and one class in the control condition. In the experimental condition, peer assessment has been implemented by using the CRT ($N = 33$). In the control condition, peer assessment has been implemented on the traditional way by using score cards ($N = 36$). This way, pupils only experienced one of the

What if pupils can assess their peers anonymously?

conditions (between-subject manipulation), but the teachers experienced both conditions (within-subject manipulation).

<Insert Figure 1 about here>

All steps of the research procedure are depicted in Figure 1. Following the suggestions of Caldwell (2007) and Trees & Jackson (2007), one of the researchers first introduced the participants to the goals of this study, the theoretical background of peer assessment and the use of CRT or score cards, depending on the condition they were assigned to. After this introduction the teacher gave a detailed explanation about the goals of the group assignment. Subsequently, pupils got an intensive training which is seen as a necessity to guarantee successful peer assessment (Dochy, Segers, & Sluijsmans, 1999; Sluijsmans et al., 2004; Tsivitanidou, Zacharia, & Hovardas, 2011). The training started with a class brainstorm about which criteria might be important when assessing a (group) presentation, followed by an exercise in which pupils were asked to think about observable behavior for every possible score (1-5) on a certain criterion. Yet, after activating pupils' prior knowledge about possible criteria during the brainstorm, the third author reflected upon these criteria and completed the rubric based on an rubric used in previous research of Raes et al. (2013) and the required standards in Flemish secondary education. The final rubric included five criteria, that is quality of content, body language, eloquence, media use and creativity. An example of such a rubric can be found in Table 1. In the last part of the training, students practiced the use of these rubrics for assessment by means of an exercise in which they had to evaluate a presentation that was shown to them in video-format.

<Insert table 1 about here>

In the courses that followed the training pupils gave their presentations, which were scored by means of peer assessment by using CRT in the experimental condition and by using

What if pupils can assess their peers anonymously?

score cards in the control condition. The presentation of every group was assessed based on the five criteria described above, using the corresponding rubrics.

In the experimental condition, the results of the peer assessment were automatically summarized and projected on a screen in front of the class by means of the CRT transmitters. No one knew which scores were given by whom. Pupils in the control condition had to give their scores by raising a scorecard, meaning that the scores were visible and non-anonymous. In both conditions the teachers also provided a score, either anonymously by using CRT or visible by raising a score card, corresponding the condition of the pupils. Finally all the group presentations were also scored by the third author who assisted the peer assessment interventions from the start.

Since the assessment aimed to have a formative rather than summative nature, the assessment intervention in both conditions ended with an additional oral feedback based on the given scores and moderated by the teacher. It has been found that such a feedback moment in which all groups were provided with some general strengths and weaknesses to take into account in the future, is a valuable extension of the scoring process during the assessment procedure (Hattie, 2003; Raes et al., 2013).

Finally, all pupils and both teachers were asked to complete an individual questionnaire to gauge pupils' and teachers' experiences with the peer assessment in either the CRT condition or the card raising condition. In addition, in the four classrooms a class discussion was organized, moderated by one of the authors in which general thoughts about the peer assessment intervention could be expressed.

3.3 Measures

This study used a mixed-method approach to obtain a complete picture of the interventions and its effects, thereby overcoming weaknesses and biases of single approaches (Denscombe, 2008).

What if pupils can assess their peers anonymously?

The quantitative data consisted of the results on the individual questionnaire filled in by all students in both conditions after the intervention. The questionnaire contained items on a 5-point Likert scale (1= totally disagree, 5= totally agree), measuring different variables which are depicted in Table 2. As indicated in Table 2, the scales and items are mostly based on previous research of Raes et al. (2013). First, the experienced anonymity and the perceived rating fairness of the peer assessment form was measured. Second, pupils' overall attitudes towards peer assessment as they experienced it were measured. Third, to measure the undesirable social effects that might accompany the peer assessment procedure, items were added that measured the general experienced peer pressure and the experienced fear of disapproval associated to giving a negative evaluation. The latter has been based on the Performance Failure Appraisal Inventory of Conroy (2002). Furthermore, feelings of discomfort when receiving face-to-face feedback were measured.

<Insert table 2 about here>

Next to quantitative data, different forms of qualitative data were gathered. The qualitative data were obtained through complete observations of the interventions registered in a logbook, a final class discussion in the four participating classes, a final remark question in the questionnaire, and a semi-structured interview with both teachers with open questions gauging their experiences with both forms of peer assessment (e.g. "Do you see a general difference between the two forms of peer assessment?", "Which form has an added value?", and "Which form do you prefer?"). Moreover, at the end of the questionnaire, students were asked in an open question to express their thoughts and remarks about the peer assessment practice they experienced. These data were used to add nuance and contour to the study, enriching it beyond what quantitative analysis can offer.

What if pupils can assess their peers anonymously?

4. Results

4.1 Manipulation check for experienced anonymity

Before the main results of this study will be presented, it is vital to verify whether the intervention has been experienced as was planned. Since it is questioned whether there is a difference between an anonymous and a non-anonymous peer assessment setting with regard to some interpersonal variables, a manipulation check was conducted to check if the CRT-condition was indeed experienced as more anonymous compared to the score card raising condition.

An ANOVA analysis with the anonymity-scale as a dependent variable showed that the experimental condition in which pupils used CRT for peer assessment was experienced as more anonymous ($M = 3.61$) than the control condition in which pupils used cards for peer assessment ($M = 2.05$, $F(1,63) = 62.51$, $p < .001$). We can conclude that our manipulation was successful.

4.2 RQ1: What is the effect of providing anonymity during peer assessment on undesirable social effects?

The answer to this question is twofold. First, concerning the perspective of the assessor (hypothesis 1a), the aim was to verify that anonymous peer assessment decreases negative social effects that is pupils' experienced peer pressure and fear of disapproval when evaluating. Second, from an assessee's point of view, it was hypothesized that students can feel uncomfortable when receiving immediate visual feedback in front of the class.

Regarding the first hypothesis, a MANOVA analysis with the two scales that measure undesirable effects when assessing as dependent variables (i.e., experienced peer pressure and fear of disapproval towards others) showed a significant effect of the peer assessment condition ($F(2,62) = 4.75$, $p < .05$), so the hypothesis could be confirmed. As shown in Figure 2, univariate analyses showed that pupils experienced significantly less peer pressure

What if pupils can assess their peers anonymously?

in the anonymous CRT condition ($M = 2.30$) compared to the non-anonymous card-raising condition ($M = 2.87$) ($F(1,63) = 5.31, p < .05$) and significantly less fear of disapproval ($F(1,63) = 6.91, p < .05$) (respectively $M = 2.10$ and $M = 2.59$).

These results have been elaborated upon by analyzing qualitative data. The hypothesis that increased anonymity is associated with decreased peer pressure is also supported by the statements of the pupils and the teachers:

'With clickers you give marks that you believe to be correct, when raising a score card you'll be influenced by peer pressure' (Pupil B., CRT-condition).

'In the class without turning point, they didn't dare to give a fail because it was visible for everyone' (Teacher A).

Regarding the hypothesis 1b concerning the perspective of the assessee, an ANOVA-analysis was conducted with the condition of peer assessment as a factor and the scale that measured the feelings of discomfort when evaluated as a dependent variable. As shown in figure 2 in both conditions students did not indicate feelings of discomfort ($M = 2.51$ regarding the CRT-condition and $M = 2.64$ regarding the card-raising condition). Moreover no significant difference has been found between the conditions ($F(1,63) = .40, p = .53$).

<Insert figure 2 about here>

4.3 RQ2: What is the effect of providing anonymity on pupils' evaluations of peer assessment?

Second it was hypothesized that pupils will have more positive attitudes towards anonymous peer assessment due to the decrease of negative social effects which has been confirmed previously. An ANOVA analysis with the attitude-scale as a dependent variable showed that pupils feel significantly more positive about peer assessment when they use CRT ($M = 4.05$), than when they have to raise cards ($M = 3.63$) ($F(1,63) = 4.51, p < .05$) and so the third hypothesis can be confirmed. This finding has also been elaborated upon by

What if pupils can assess their peers anonymously?

analyzing the qualitative data and the result is supported by the qualitative data, as 64% of the pupils who gave a final remark in the questionnaire explicitly reported positive feelings towards the anonymous assessment using CRT:

'I liked scoring the work of my peers, particularly with the use of CRT' (Pupil M., CRT-condition).

'I had real fun working with the CRT. We should definitely do this again' (Pupil H., CRT-condition).

4.4 RQ 3: Is the anonymous peer assessment procedure using CRT a valid method for assessment ?

Given the lower scores on peer pressure and fear of disapproval and the more positive attitudes and feelings of comfort in the anonymous condition, the third hypothesis was that the anonymous peer assessment procedure would be more valid compared to the traditional peer assessment procedure. To measure and compare the validity of both peer assessment interventions in secondary education, the *criterion related validity* was tested by calculating the correlation between pupils' scores and teachers' scores in correspondence to Falchikov & Goldfinch (2000) and Sung, Chang, Chang, and Yu (2010). In the control condition where the pupils had to raise cards, the correlation between the pupils' scores and the teachers' scores was 0.93, which is exceptionally high (Falchikov & Goldfinch, 2000; Sung et al., 2010). In the experimental condition where pupils gave their score by means of CRT, the correlation was only .42, which is rather low. This seems to indicate that raising cards is a more valid way of assessment than using CRT, which is contrary to our hypothesis.

However, to verify these results, *face validity* has been measured as well. To measure the face validity as judged by the pupils, scores on perceived rating fairness were analyzed. An ANOVA analysis with the perceived rating fairness as dependent variable showed that

What if pupils can assess their peers anonymously?

pupils perceived both forms of peer assessment as rather fair ($M = 3.50$ for raising cards and $M = 3.10$ for CRT, $F(1,63) = 3.97, p > .05$).

However, four pupils reported their doubt about the fairness of the assessment with CRT:

'I had the feeling that not everyone was giving scores in a fair way.' (Pupil S., CRT-condition)

Some students and one of the teachers gave a possible explanation for this doubt, stating:

'I was happy that we had to raise cards, otherwise the given scores were way too low.' (Pupil V.M., card raising-condition)

'The class with CRT received lower scores, the class without CRT received higher scores. The pupils liked CRT until they saw the consequence: lower scores'.
(Teacher A)

The pupils indeed seemed to give lower scores in the CRT-condition ($\beta = -0.35, F(1,67) = 9.89; p < .05$). However, this does not mean that the assessment was not valid. It seems plausible that because of the reduction in peer pressure and fear of disapproval, pupils dared to give lower and more correct scores. This is confirmed by the teachers, as they reported that they believe that the students are capable of evaluating each other after the training and when using rubrics. Also, they reported that the evaluation happened correctly and objectively, as becomes apparent in the following statement:

'I am sure that pupils can evaluate each other objectively and correctly, if it is a sum of different scores on different facets, as it was done' (Teacher B).

What if pupils can assess their peers anonymously?

Moreover, they indeed report that the assessment is only valid when peer pressure is reduced:

‘...pupils can evaluate each other, but only if you can sufficiently reduce the peer pressure’ (Teacher A).

Next to the pupils and the teachers, the observing researcher (i.e., the third author) reported to believe that the assessment process was valid, especially in the CRT-condition. She reported that in the non-anonymous card-raising condition, the teachers sometimes looked at the scores of the pupils before choosing the card they would be raising. This can explain the exceptionally high correlation between the teacher’s scores and the pupils’ scores in the non-anonymous condition and therefore also the contradiction between the *face validity* measures that confirm our hypothesis (as judged by the pupils, the teachers and the observing researcher) and the *criterion related validity* as described above.

4.5 RQ4: What is the effect of providing anonymity during peer assessment on teachers’ experiences of peer assessment?

To answer this research question, the responses of the two teachers during the semi-structured interview were analyzed. Both teachers indicated that peer assessment had an added value over traditional assessment forms. They confirm the important advantages of peer assessment, as becomes apparent in the following statement:

‘It increases self-awareness, and it also teaches them to reinterpret the meaning of the grading system’ (Teacher B).

They both mentioned the assessment with CRT as the preferable choice, because of the anonymity of the assessment:

What if pupils can assess their peers anonymously?

'The anonymity of CRT is an extra surplus because it takes away the pressure'

(Teacher B).

However, some of their comments pointed towards important considerations in this process. First of all, teacher A emphasizes the importance of the control of the teacher. She points out that, while peer pressure is reduced in the CRT-condition, pupils could still give their friends higher scores, because no one would know. Therefore, while she prefers CRT over card-raising, she feels that this new technology can only work if there is enough control of the teacher, for example by making the scores anonymous towards other pupils but not towards the teacher.

Secondly, teacher B emphasizes the importance of class characteristics. She reports:

'The stress in both of my classes was the same: one class had the advantage of voting anonymously, the other class was more 'mature' for the card-system'.

Hereby she refers to class characteristics like the age and maturity of the pupils in the classes. She stresses that the class that was now randomly assigned to the CRT-condition, would have had much more problems with the card-raising system. She mentions that the anonymity of assessing with CRT is especially important with younger teenagers and teenagers that are not mature enough to cope with peer pressure.

5. Discussion

This research starts from the fact that peer assessment has proven to be a promising assessment form (see section 1.1), but that it is often influenced by undesirable social effects (Van den Bossche, Gijsselaers, Segers, & Kirschner, 2006; van Gennip et al., 2010). First, negative feelings during peer assessment might undermine the motivation for participation (Stepanyan et al., 2009) and second, reciprocity effects potentially jeopardize the validity of the assessment procedure (Strijbos & Sluijsmans, 2010).

What if pupils can assess their peers anonymously?

Following different social impact theories (e.g., Deutsch & Gerard, 1955; Latané, 1981), anonymity has been proposed to counter these effects. In a face-to-face classroom setting, anonymous peer assessment can be established by means of CRT (Raes et al., 2013), yet research about the effects of anonymous peer assessment is lacking (Panadero et al., 2013), particularly in secondary education. In this regard, a quasi-experimental setting has been set up to verify whether an anonymous peer assessment setting could indeed reduce peer pressure and feelings of discomfort (RQ 1), consequently induce more positive attitudes towards peer assessment (RQ2), and thereby add to the validity of the peer assessment procedure (RQ3).

Based on quantitative as well as qualitative analyses the results confirm the first hypothesis that pupils experienced less peer pressure and less fear of disapproval when they could give their scores anonymously using CRT, compared to when they had to give their scores non-anonymously by raising cards (hypotheses 1a). Moreover, this led to the fact that students felt more positive towards peer assessment, which means that also hypothesis two could be confirmed. These findings are in line with previous research in higher education (Howard et al., 2010; Lu & Bol, 2007; Raes et al., 2013). Moreover, hypothesis 1b taking into account the assessee's point of view has not been confirmed meaning that students did not report to feel more discomfort when the summarized visual feedback was given in front of the whole class. Given these results, anonymous peer assessment using CRT in a face-to-face class setting might be preferable over non-anonymous peer assessment.

The quantitative and qualitative results obtained to confirm hypothesis three regarding the validity of the anonymous peer assessment procedure are however less straightforward. While construct validity, as measured by the correlation between the teacher's scores and the pupils' scores, contradicts the proposed hypothesis that an anonymous peer assessment

What if pupils can assess their peers anonymously?

procedure is more valid, face validity confirms this hypothesis. As a result, the statement that anonymous peer assessment using CRT is preferable, needs to be put in perspective.

Regarding the fourth research question, it was further interesting to see that teachers in secondary education were positive about the anonymous CRT setting as well, and do call it a valid and objective way of assessment if a good training was given and if rubrics are used. However, they emphasize the importance of teacher control in this process. This is in line with the suggestions of Hovardas et al. (2014), that guidance is needed during the whole dynamic process of peer assessment. Teacher control might prevent pupils to give higher grades to their friends, and might give even more feelings of comfort, since pupils often rely on the 'grading experience' of their teachers (Strijbos, Narciss, & Dünnebier, 2010; Yang, Badger, & Yu, 2006). In addition, Gielen, Dochy, Onghena, Janssens, and Decuyper (2007) found that students used to a teacher-led assessment can get frustrated when the teacher feedback opportunities are being totally replaced by peer feedback.

Concerning the validity problem discussed above, this frustration might also explain why some students reported that the peer assessment process using CRT was not totally fair. Pupils could not compare the scores given by the peers with the scores given by the teacher, which was possible in the card raising condition. However, as stated in section 4.4, a more plausible explanation for the unfairness of the assessment according to some pupils can be found in the lower scores that were given using CRT. Given the high face validity of the CRT assessment setting, as reported by most pupils, the teachers and the observing researcher, it seems likely that although the lower scores were perceived by some pupils as unfair, these were actually fair scores. Moreover, it was found in this study that in the card raising condition, the teacher felt pressure as well and was influenced by the scores given by the pupils. On the other hand, it is also plausible that pupils may be influenced by the score given by the teacher as well. Although this has been countered by the observing researcher, future

What if pupils can assess their peers anonymously?

research should include a condition without teacher's scores to increase the comparability between the two groups and to check whether the scores given by the pupils of the two groups are the same.

In this context the often used measure of construct validity in peer assessment research, that is the correlation between the pupils' scores and the scores of the teacher, who is seen as an expert and whose score serves as 'the unbiased baseline' (Panadero et al., 2013) is questionable. Taking this into account, the face validity captured in this study might be the more adequate measurement of validity.

Some of the problems described above, such as the lack of teacher control and the corresponding frustration of pupils, can be solved with some small changes in the procedure. First, settings of the CRT can be changed so that pupils can remain anonymous for each other when giving feedback, but can be identified by the teacher afterwards. This form of teacher control might prevent pupils to give unfairly high scores to friends, while it still reduces peer pressure and fear of disapproval towards friends. Second, it is possible for the teacher to give a score non-anonymously, so that his or her score appears on the screen separately from the scores of the pupils as a visual feedback. This way the teacher's and pupils' score cannot be influenced by each other scores as might be the case in the card-raising condition, but the discomfort teenagers experience when not receiving a score from the teacher can be reduced.

Considering the current teacher-led assessment culture, we advocate that future research on peer assessment and feedback should leave the strict distinction between peer and teacher assessment behind and acknowledge the richness of the triangulation of assessment and feedback sources co-assessment practices offer. New response technologies have the potential to offer the needed immediacy of formative assessment and feedback provided by peers as well as by teachers (Topping, 2009). Future research should address this issue and

What if pupils can assess their peers anonymously?

investigate the differences in acceptance of peer and/or expert assessment and feedback (Hovardas et al., 2014).

Although the results of this research have an invaluable importance for further peer assessment research, and for practitioners who want to implement peer assessment in secondary education, some limitations of this study need to be taken into account. First of all, since only four classes participated, it was impossible to measure the effects caused by classroom characteristics. For example, Deutsch and Gerard (1955) stated that individuals that form close groups are more susceptible of peer pressure and social influences than individuals not forming a close group. The levels of peer pressure and fear of disapproval found in this study are still rather low (see Figure 2). Following Deutsch and Gerard (1955), this might be caused by specific class-group characteristics. The importance of these class characteristics has also been pointed by one of the teachers (see section 4.5). A large scale study is necessary to confirm and deepen these results, taking into account different class cultures.

Second, all pupils out of the same class were assigned to one peer assessment setting. Given this between-subject manipulation, pupils could not make a direct comparison between the two conditions themselves, thereby revealing their preferences. A replication of this study using a within-subject manipulation might give us more insight into pupils' attitudes and emotions with regard to these different forms of peer assessment.

Third, it might be stated that we are only facing a novelty effect, that is, students like CRT more since it is a new technology used in the classroom. In this case, this effect would disappear if this technology would really be used in classes on a regular basis. Draper and Brown (2004) found however that this novelty effect exists for the use of CRT in classes, but only lasts for something between 5 and 50 minutes. After this there is more of an anti-novelty effect which means that students tend to be rather skeptical. Since the intervention in this

What if pupils can assess their peers anonymously?

study took several hours, there is no reason to believe our positive effects would fade away with more time.

Finally, although anonymous peer assessment proved to be promising in this study, the goals that are put forth during the assessment need to be considered when choosing an assessment form. For example, when aiming at the development of social skills, a non-anonymous setting might still be preferred. The study of van Gennip et al. (2010) shows that the social aspects of peer assessment are not unfavorable by nature. The question is how to learn students cope with these social aspects that are inherent to the peer assessment process. In this study we only focused on how interpersonal variables influence the course of the peer assessment process and we therefore see anonymity as a catalyst that may help gauging the change in student perceptions towards peer assessment and feedback. We assume that anonymity might only be a phase in establishing effective peer assessment practices within safe learning contexts. The peer assessment practices described in this research are therefore part of the challenging process of developing sustainable assessment skills, where students become more able to exercise evaluative judgment (Boud, Lawson, & Thompson, 2013; Boud, 2007; Fastré et al., 2013). In this process, we introduce anonymity as a facilitator to make learning (and feedback) as visible as possible in the classroom, allowing students to practice while acknowledging the dispositions of students and the power of peers in the feedback process (Hattie & Gan, 2011).

6. References

- Ainsworth, S., Gelmini-Hornsby, G., Threapleton, K., Crook, C., O'Malley, C., & Buda, M. (2011). Anonymity in Classroom Voting and Debating. *Learning and Instruction*, 21(3), 365–378.
- Author (2012). Blinded for review.

What if pupils can assess their peers anonymously?

- Barron, B. (2003). When smart groups fail. *Journal of the Learning Sciences*, 12, 307–359.
doi:10.1207/S15327809jls1203_1
- Biggs, J. (1996). Enhancing teaching through constructive alignment. *Higher Education*, 32(3), 347–364. doi:10.1007/BF00138871
- Birenbaum, M. (2003). New insights into learning and teaching and their implications for assessment. In M. Segers, F. Dochy, & E. Cascallar (Eds.), *Optimising New Modes of Assessment: In Search of Qualities and Standards* (pp. 13–37). Dordrecht, The Netherlands: Kluwer Academic.
- Black, P., & Wiliam, D. (1998). Assessment and Classroom Learning. *Assessment in Education: Principles, Policy & Practice*, 5(1), 7–74. doi:10.1080/0969595980050102
- Blasco-Arcas, L., Buil, I., Hernández-Ortega, B., & Sese, F. J. (2013). Using clickers in class. The role of interactivity, active collaborative learning and engagement in learning performance. *Computers & Education*, 62, 102–110.
doi:10.1016/j.compedu.2012.10.019
- Bojinova, E., & Oigara, J. (2011). Teaching and Learning with Clickers: Are Clickers Good for Students? *Interdisciplinary Journal of E-Learning and Learning Objects*, 7(1), 169–184.
- Boud, D. (2007). Reframing assessment as if learning were important. *Rethinking Assessment in Higher Education: Learning for the Longer Term*, 14–25.
- Boud, D., Cohen, R., & Sampson, J. (1999). Peer Learning and Assessment. *Assessment & Evaluation in Higher Education*, 24(4), 413–426.
- Boud, D., Lawson, R., & Thompson, D. G. (2013). Does student engagement in self-assessment calibrate their judgement over time? *Assessment & Evaluation in Higher Education*, 1–16.

What if pupils can assess their peers anonymously?

- Brady, M., Seli, H., & Rosenthal, J. (2013). "Clickers" and metacognition: A quasi-experimental comparative study about metacognitive self-regulation and use of electronic feedback devices. *Computers and Education*, 65, 56–63.
- Caldwell, J. E. (2007). Clickers in the Large Classroom: Current Research and Best-Practice Tips. *CBE— Life Sciences Education*, 6(1), 9–20. doi:10.1187/cbe.06-12-0205
- Cartney, P. (2010). Exploring the use of peer assessment as a vehicle for closing the gap between feedback given and feedback used. *Assessment & Evaluation in Higher Education*, 35(5), 551–564. doi:10.1080/02602931003632381
- Conroy, D. E. (2002). *The Performance Failure Appraisal Inventory. User's Manual*. The Pennsylvania State University, department of Kinesiology. Retrieved from http://www.personal.psu.edu/dec9/uploads/3/0/4/0/304067/2003_pfai_users_manual.pdf
- De Grez, L., Valcke, M., & Roozen, I. (2009). The impact of an innovative instructional intervention on the acquisition of oral presentation skills in higher education. *Computers & Education*, 53(1), 112–120. doi:10.1016/j.compedu.2009.01.005
- Deutsch, M., & Gerard, H. B. (1955). A study of normative and informational social influences upon individual judgment. *The Journal of Abnormal and Social Psychology*, 51(3), 629–636. doi:10.1037/h0046408
- Denscombe, M. (2008). Communities of Practice: A Research Paradigm for the Mixed Methods Approach. *Journal of Mixed Methods Research*. doi:10.1177/1558689808316807
- Dochy, F., Segers, M., & Sluijsmans, D. (1999). The use of self-, peer and co-assessment in higher education: A review. *Studies in Higher Education*, 24(3), 331–350. doi:10.1080/03075079912331379935

What if pupils can assess their peers anonymously?

- Draper, S. W., & Brown, M. I. (2004). Increasing interactivity in lectures using an electronic voting system. *Journal of Computer Assisted Learning*, 20(2), 81–94.
doi:10.1111/j.1365-2729.2004.00074.x
- Dumont, H., & Istance, D. (2010). Analysing and designing learning environments for the 21st century. In H. Dumont, D. Instance, & F. Benavides (Eds.), *The Nature of Learning: Using Research to Inspire Practice* (pp. 19–34). Organisation for Economic Co-operation and Development. Retrieved from <http://www.oecd-ilibrary.org/content/chapter/9789264086487-3-en>
- Falchikov, N., & Goldfinch, J. (2000). Student Peer Assessment in Higher Education: A Meta-Analysis Comparing Peer and Teacher Marks. *Review of Educational Research*, 70(3), 287–322. doi:10.3102/00346543070003287
- Fastré, G. M. J., van der Klink, M. R., Sluijsmans, D., & van Merriënboer, J. J. G. (2013). Towards an integrated model for developing sustainable assessment skills. *Assessment & Evaluation in Higher Education*, 38(5), 611–630.
doi:10.1080/02602938.2012.674484
- Gielen, D., Dochy, F., Onghena, P., Janssens, S., & Decuyper, S. (2007). A complementary role for peer feedback and staff feedback in powerful learning environments. In D. Gielen (Ed.), *Peer assessment as a tool for learning* (pp. 157–199). Leuven: Katholieke universiteit Leuven, Faculteit Psychologie en Pedagogische Wetenschappen Centrum voor Opleidingsdidactiek.
- Gielen, M., & De Wever, B. (2012). Peer Assessment in a Wiki: Product Improvement, Students' Learning And Perception Regarding Peer Feedback. *Procedia - Social and Behavioral Sciences*, 69, 585–594. doi:10.1016/j.sbspro.2012.11.450
- Han, J. H., & Finkelstein, A. (2013). Understanding the effects of professors' pedagogical development with Clicker Assessment and Feedback technologies and the impact on

What if pupils can assess their peers anonymously?

students' engagement and learning in higher education. *Computers and Education*, 65, 64–76.

Harris, L. R., & Brown, G. T. L. (2013). Opportunities and obstacles to consider when using peer- and self-assessment to improve student learning: Case studies into teachers' implementation. *Teaching and Teacher Education*, 36, 101–111.
doi:10.1016/j.tate.2013.07.008

Hattie, J. (2003). *Formative and summative interpretations of assessment information*. Retrieved from
<http://www.education.auckland.ac.nz/webdav/site/education/shared/hattie/docs/formative-and-summative-assessment-%282003%29.pdf>

Hattie, J., & Gan, M. (2011). Instruction based on feedback. In R. E. Mayer & P. A. Alexander (Eds.), *Handbook of research on learning and instruction* (pp. 249–271). New York.

Hovardas, T., Tsivitanidou, O. E., & Zacharia, Z. C. (2014). Peer versus expert feedback: An investigation of the quality of peer feedback among secondary school students. *Computers & Education*, 71, 133–152. doi:10.1016/j.compedu.2013.09.019

Howard, C. D., Barrett, A. F., & Frick, T. W. (2010). Anonymity to Promote Peer Feedback: Pre-Service Teachers' Comments in Asynchronous Computer-Mediated Communication. *Journal of Educational Computing Research*, 43(1), 89–112.
doi:10.2190/EC.43.1.f

Kay, R. H., & LeSage, A. (2009). Examining the benefits and challenges of using audience response systems: A review of the literature. *Computers & Education*, 53(3), 819–827.
doi:10.1016/j.compedu.2009.05.001

What if pupils can assess their peers anonymously?

- Kollar, I., & Fischer, F. (2010). Peer assessment as collaborative learning: A cognitive perspective. *Learning and Instruction*, 20(4), 344–348.
doi:10.1016/j.learninstruc.2009.08.005
- Latané, B. (1981). The psychology of social impact. *American Psychologist*, 36(4), 343–356.
doi:10.1037/0003-066X.36.4.343
- Li, L., Liu, X., & Steckelberg, A. L. (2010). Assessor or assessee: How student learning improves by giving and receiving peer feedback. *British Journal of Educational Technology*, 41(3), 525–536. doi:10.1111/j.1467-8535.2009.00968.x
- Liu, X., & Li, L. (2013). Assessment training effects on student assessment skills and task performance in a technology-facilitated peer assessment. *Assessment & Evaluation in Higher Education*, 0(0), 1–18. doi:10.1080/02602938.2013.823540
- Lu, R., & Bol, L. (2007). A comparison of anonymous versus identifiable e-peer review on college student writing performance and the extent of critical feedback. *Journal of Interactive Online Learning*, 6(2), 100–115.
- Panadero, E., Romero, M., & Strijbos, J.-W. (2013). The impact of a rubric and friendship on peer assessment: Effects on construct validity, performance, and perceptions of fairness and comfort. *Studies in Educational Evaluation*.
doi:10.1016/j.stueduc.2013.10.005
- Pope, N. K. L. (2005). The impact of stress in self- and peer assessment. *Assessment & Evaluation in Higher Education*, 30(1), 51–63. doi:10.1080/0260293042003243896
- Raes, A., Vanderhoven, E., & Schellens, T. (2013). Increasing anonymity in peer assessment by using classroom response technology within face-to-face higher education. *Studies in Higher Education*, 0(0), 1–16. doi:10.1080/03075079.2013.823930

What if pupils can assess their peers anonymously?

- Scardamalia, M. (2001). Big Change Questions Will Educational Institutions, within their Present Structures, be Able to Adapt Sufficiently to Meet the Needs of the Information Age? *Journal of Educational Change*, 2(2), 171–176. doi:10.1023/A:1017982102060
- Sluijsmans, D. M. A. (2002). *Student involvement in assessment: The training of peer assessment skills. (Unpublished doctoral dissertation)*. Open University of the Netherlands, Heerlen.
- Sluijsmans, D. M. A., Brand-Gruwel, S., van Merriënboer, J. J. G., & Martens, R. L. (2004). Training teachers in peer-assessment skills: effects on performance and perceptions. *Innovations in Education and Teaching International*, 41(1), 59–78. doi:10.1080/1470329032000172720
- Smith, H., Cooper, A., & Lancaster, L. (2002). Improving the Quality of Undergraduate Peer Assessment: A Case for Student and Staff Development. *Innovations in Education and Teaching International*, 39(1), 71–81. doi:10.1080/13558000110102904
- Steinberg, L., & Monahan, K. C. (2007). Age Differences in Resistance to Peer Influence. *Developmental Psychology*, 43(6), 1531–1543. doi:10.1037/0012-1649.43.6.1531
- Stepanyan, K., Mather, R., Jones, H., & Lusuardi, C. (2009). Student Engagement with Peer Assessment: A Review of Pedagogical Design and Technologies. In M. Spaniol, Q. Li, R. Klamma, & R. W. H. Lau (Eds.), *Advances in Web Based Learning – ICWL 2009* (pp. 367–375). Springer Berlin Heidelberg. Retrieved from http://link.springer.com/chapter/10.1007/978-3-642-03426-8_44
- Strijbos, J.-W., Narciss, S., & Dünnebier, K. (2010). Peer feedback content and sender's competence level in academic writing revision tasks: are they critical for feedback perceptions and efficiency? *Learning and Instruction*, 20, 291–303.

What if pupils can assess their peers anonymously?

- Strijbos, J.-W., & Sluijsmans, D. (2010). Unravelling peer assessment: Methodological, functional, and conceptual developments. *Learning and Instruction*, 20(4), 265–269. doi:10.1016/j.learninstruc.2009.08.002
- Sumter, S. R., Bokhorst, C. L., Steinberg, L., & Westenberg, P. M. (2009). The developmental pattern of resistance to peer influence in adolescence: will the teenager ever be able to resist? *Journal of Adolescence*, 32(4), 1009–1021. doi:10.1016/j.adolescence.2008.08.010
- Sung, Y.-T., Chang, K.-E., Chang, T.-H., & Yu, W.-C. (2010). How many heads are better than one? The reliability and validity of teenagers' self- and peer assessments. *Journal of Adolescence*, 33(1), 135–145. doi:10.1016/j.adolescence.2009.04.004
- Topping, K. J. (2009). Peer Assessment. *Theory Into Practice*, 48, 20-27. doi:10.1080/00405840822577569
- Topping, K. J. (1998). Peer Assessment between Students in Colleges and Universities. *Review of Educational Research*, 68(3), 249. doi:10.2307/1170598
- Topping, K. J. (2003). Self and peer assessment in school and university: Reliability, validity and utility. In M. Segers, F. Dochy, & E. Cascallar (Eds.), *Optimising New Modes of Assessment: In Search of Qualities and Standards* (pp. 55–87). Dordrecht, The Netherlands: Kluwer Academic.
- Topping, K. J. (2010). Methodological quandaries in studying process and outcomes in peer assessment. *Learning and Instruction*, 20, 339–343.
- Trees, A. R., & Jackson, M. H. (2007). The learning environment in clicker classrooms: student processes of learning and involvement in large university-level courses using student response systems. *Learning, Media and Technology*, 32(1), 21–40. doi:10.1080/17439880601141179

What if pupils can assess their peers anonymously?

- Tsivitanidou, O. E., Zacharia, Z. C., & Hovardas, T. (2011). Investigating secondary school students' unmediated peer assessment skills. *Learning and Instruction*, 21(4), 506–519. doi:10.1016/j.learninstruc.2010.08.002
- Van den Bossche, P., Gijssels, W. H., Segers, M., & Kirschner, P. A. (2006). Social and Cognitive Factors Driving Teamwork in Collaborative Learning Environments Team Learning Beliefs and Behaviors. *Small Group Research*, 37(5), 490–521. doi:10.1177/1046496406292938
- Van Gennip, N. A. E., Segers, M. S. R., & Tillema, H. H. (2009). Peer assessment for learning from a social perspective: The influence of interpersonal variables and structural features. *Educational Research Review*, 4, 41–54. doi:10.1016/j.edurev.2008.11.002
- Van Gennip, N. A. E., Segers, M. S. R., & Tillema, H. H. (2010). Peer assessment as a collaborative learning activity: The role of interpersonal variables and conceptions. *Learning and Instruction*, 20(4), 280–290. doi:10.1016/j.learninstruc.2009.08.010
- Van Zundert, M., Sluijsmans, D., & van Merriënboer, J. (2010). Effective peer assessment processes: Research findings and future directions. *Learning and Instruction*, 20(4), 270–279. doi:10.1016/j.learninstruc.2009.08.004
- Yang, M., Badger, R., & Yu, Z. (2006). A comparative study of peer and teacher feedback in a Chinese EFL writing class. *Journal of Second Language Writing*, 15(3), 179–200. doi:10.1016/j.jslw.2006.09.004